3J-3

2パス限定投機システムにおける投機コードの最適化

本間 勇貴 † 十鳥 弘泰 † 大津 金光 † 大川 猛 † 横田 隆史 †
† 宇都宮大学大学院工学研究科

1 はじめに

我々はスレッドレベル並列性を抽出する投機的マルチスレッド処理方式に基づくマルチコアプロセッサアーキテクチャである 2 パス限定投機システム PALS[1] を開発している．本稿では，本システムで使用する投機スレッドコードの最適化手法を検討し，その適用効果を評価する．最適化技術として，PALS で使用する投機スレッドコードに対してループ展開とリストスケジューリングを適用する手法を検討し評価する．

2 2パス限定投機方式

2 パス限定投機システム PALS では，まず，パスプロファイリングにより投機対象のループ中に存在する実行経路（パス）の中で最も実行頻度の高いパス（#1 パス）と 2 番目に実行頻度の高いパス（#2 パス）の情報を得ておき，それらの情報を基に，それぞれのパスに特化した投機スレッドコードと元のループ構造を保持した非投機スレッドコードをあらかじめ用意する．マルチスレッド実行時には#1 パスと#2 パスのどちらを実行するべきか予測しながら投機的にマルチスレッド処理する．投機失敗時にはスレッドの実行を破棄し，回復処理を行った後，別のスレッドコードを実行する．1 回目の投機失敗時には予測したパスとは違うパスを実行し，2 回目の投機失敗時には非投機スレッドコードを実行する．

PALS には，ループのイテレーションごとにスレッドの生成や終了などのスレッド制御やレジスタによるスレッド間の高速通信などの機能がある．スレッド制御やスレッド間通信が完了するまでの待ち時間は，マルチスレッド実行を行う上でオーバーヘッドとなる．このオーバーヘッドを削減することで，プログラム実行の高速化が期待できる．

3 ループ展開の適用

PALS におけるスレッド制御のオーバーヘッドの削減にはループ展開が有効である [2]．PALS 向けの投機スレッドコードにループ展開を適用することで，スレッドサイズが増加し，スレッド制御のオーバーヘッドを相対的に小さくすることができる．また，ループ展開によりイテレーション数が減少するため，スレッド制御をする回数の削減も可能である．

ループ展開において，ループ展開数が性能上最も重要なパラメータである．ループ展開を適用することでパスの総数が増え，実行頻度上位 2 本のパスの実行割合が減少する．ループ展開でスレッドサイズを増やしたことにより性能向上する一方で，ループ展開により実行頻度上位 2 パスの実行割合が減少し，パス予測失敗率の増加することにより性能低下の可能性がある．そこで本研究では，PALS 向けの投機スレッドコードにループ展開を適用する際には，計算モデルを用いた解析的な投機実行サイクル数の見積りにより，最適と予測されるループ展開数を決定する．以下の制約から，投機実行サイクル数の見積りの範囲を決定することで，投機実行サイクル数の見積りのコストを最小限に抑える．

- 命令キャッシュのサイズ
- 1 スレッドあたりの実行サイクル数
- 投機的ストアアクセス数

4 マルチスレッドコードのスケジューリング

スレッド間においてスレッドコード内のスレッド間依存変数を最初に使用する命令（値使用命令）とスレッド間依存変数を最後に定義する命令（値定義命令）でレジスタを介したスレッド間通信を行う．値使用命令を実行するためには前のスレッドで確定したスレッド間依存変数の値を受信する必要がある．また，値使用命令がスレッド間依存変数の値を受け取るまでの待ち時間は，値使用命令と値定義命令の間に存在する命令数により変化する．PALS の投機スレッドコードにループ展開を適用すると，値使用命令と値定義命令の間に存在する命令数が増加する．よって，ループ展開によりスレッド間依存変数による転送待ち時間が増大する．本稿ではスレッド間依存変数による転送待ち時間を削減する PALS 向けスレッドコードのスケジューリング手法を検討する．

PALS 向けの投機スレッドコードはパスに特化したコードになっているため，分岐命令を含まない 1 つの基本ブロックとして扱うことができる．これを踏まえて，本稿ではリストスケジューリングをベースとして PALS 向けの投機スレッドコードをスケジューリングする．リストスケジューリングを行う際の，スケジュールする命令の優先順位を以下のようにする．値使用命

(1) 値定義命令
(2) クリティカルパス上の命令
(3) 投機の成否を判定する命令
(4) 1,2,3,5,6 以外の命令
(5) メモリアクセス命令
(6) 値使用命令

Optimization of Speculative Code for Two-Path Limited Speculation System
†Yuki Homma, Hiroyoshi Jutori, Kanemitsu Ootsu, Takeshi Ohkawa and Takashi Yokota
Graduate School of Engineering, Utsunomiya University (†)

令をスレッドの後方へ移動し，値定義命令をスレッドの前方に移動することで，値使用命令と値定義命令の間にある命令数が減少する．これにより，スレッド間依存変数の転送待ち時間を削減することができる．また，クリティカルパス上の命令や投機の成否を判定する命令も優先してスケジューリングすることで，スレッド自体の実行時間や投機失敗によるコストを削減することができる．さらに，メモリアクセス命令を後方に移動することで，投機的なロードにより要求した値を受け取るまでの待ち時間の削減を図る．

5 評価

PALS向けのスレッドコードにループ展開とリストスケジューリングを適用したときの速度向上率について評価する．評価にはPALSシミュレータを用いる．評価に用いたシミュレーションパラメータを表1に示す．

SPEC CINT2000 ベンチマークプログラムのうち164.gzip, 181.mcf, 300.twolfに含まれるループの中から投機対象とするループを1つずつ選定し，それらを評価対象とした．投機実行サイクル数見積りの結果，ループ展開数は164.gzipで5，181.mcfで6，300.twolfで4となった．これらの展開数でループ展開を行い，シミュレータで実行した結果とループ展開を適用した投機スレッドコードにリストスケジューリングを適用し，シミュレータで実行した結果を図1に示す．速度向上率は最適化を適用する前のスレッドコードを実行したときのマルチスレッド実行サイクル数をそれぞれの最適化を適用したスレッドコードを実行したときのマルチスレッド実行サイクル数で割ることで算出した．

ループ展開を適用した結果，164.gzip で 1.39 倍，181.mcfで1.83倍，300.twolfで1.23倍の速度向上を達成した．181.mcfは3つの評価対象の中で最もループ展開による効果が高かったが，これは，181.mcfの投機対象ループのイテレーション数が多く，パスの出現パターンが規則的であったためである．パスの出現パターンが規則的であるとき，その規則の周期に合わせた展開数でループ展開を適用することで，ループ展開後のパスの総数が少なくすることができる．その結果，実行頻度上位2パスの実行頻度の低下を抑えることができるため，実行頻度上位2パスの高い実行頻度を維持することができ，PALSによる実行の高い効果が期待できる．

ループ展開とリストスケジューリングを適用した結果，164.gzipで1.45倍，181.mcfで2.23倍，300.twolfで1.23倍の速度向上を達成した．この場合でも181.mcfにおける速度向上率が最も高い数値となった．これは，スレッド内における値使用命令と値定義命令の間に存在する命令数が非常に多く，値定義命令と依存関係にある命令数が少なかったために，スケジューリングによる高い効果が得られた．164.gzipや300.twolfは181.mcfに比べると最適化の効果は小さかったが，これは，ループのイテレーション数が少ないことやパス予測の成功率が低いためである．このようなプログラムにおいて

表 1: シミュレーションパラメータ

スレッド制御機構	台数 4 スレッド制御にかかるサイクル数 8 サイクル 2 レベルパス予測 パスの履歴レジスタのビット長 4 カウンタテーブルのビット長 2
スレッド実行機構	4 命令同時実行アウトオブオーダスーパースカラ 5 段パイプライン
投機メモリアクセスユニット	各スレッドが実行可能なストア命令数 32 命令 レイテンシ 2 サイクル
1 次キャッシュ	4-ウェイセットアソシアティブ・LRU 方式 命令・データ 各 16K バイト レイテンシ 2 サイクル
メインメモリ	レイテンシ 100 サイクル

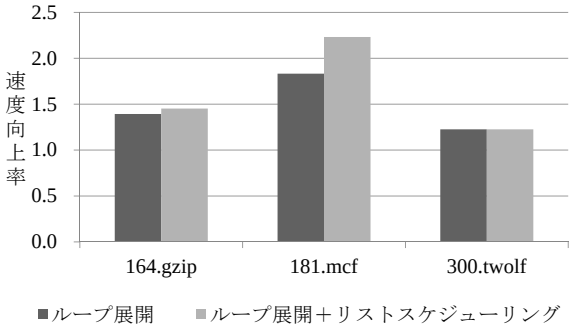

図 1: ループ展開とリストスケジューリングの効果

も，高い性能向上率を達成するためには，パス予測失敗によるコストや投機失敗自体を削減するに手段ついて検討する必要がある．

6 おわりに

本稿では2パス限定投機システムで実行する投機スレッドコードに対して，ループ展開およびリストスケジューリングを適用する手法について検討し評価した．評価の結果，2つの最適化を適用する前と比べて，最大2.23倍の速度向上を確認した．

今後の課題としては，投機失敗によるコストまたは，投機失敗回数自体を削減する手段について検討し，PALSにより多くのプログラムを性能向上できるようにすることが挙げられる．

謝辞

本研究は，一部日本学術振興会科学研究費補助金（基盤研究 (C)24500055，同 (C)24500054，同 (C)25330055，若手研究 (B)25730026）の援助による．

参考文献

[1] 十鳥 弘泰, 大津 金光, 横田 隆史, 馬場 敬信: “2パス限定投機方式を実現するマルチコアプロセッサ PALSの提案”，信学技報，Vol.109，No.319(CPSY2009-46)，pp.19-24，2009.

[2] 本間 勇貴, 十鳥 弘泰, 大津 金光, 大川 猛, 横田 隆史, 馬場 敬信: “2パス限定投機方式におけるループ展開の効果”，信学技報，Vol.112，No.173(CPSY2012-9)，pp.1-6，2012.